The directional resolution needed when measuring head-related transfer functions

Plogsties, Jan; Minnaar, Pauli; Christensen, Flemming; Olesen, Søren Krarup; Møller, Henrik

Publication date:
2000

The directional resolution needed when measuring head-related transfer functions

Jan Plogsties, Pauli Minnaar, Flemming Christensen, Søren Krarup Olesen, Henrik Møller

Department of Acoustics, Aalborg University, Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark
acoustics@acoustics.auc.dk

DAGA 2000, Oldenburg, Germany, 20 - 24 March 2000

A listening experiment was done to determine the lowest resolution with which head-related transfer function (HRTF) measurements have to be made to ensure that interpolations between them are indistinguishable from measured HRTFs. Measured and interpolated HRTFs were compared in a three-alternative forced choice test for both static and dynamic sound sources. A criterion was found that predicts the listening test results for static sound sources, which was used to determine the resolution needed when measuring HRTFs for all directions on the sphere around the head.

1 Introduction

In binaural synthesis head-related transfer functions (HRTFs) are used to implement virtual sound sources with locations relative to a listener. HRTFs are typically measured on a human or an artificial head with a limited (often low) directional resolution. In order to represent all directions around the listener interpolation is required between the measured HRTFs. The interpolation may be done either off-line (to create an HRTF database for binaural synthesis) or in real-time (called commutation).

The work presented here is a continuation of a previous study [1] that addresses the question: What is the lowest directional resolution needed when measuring HRTFs that enables the construction of the HRTFs of all directions? Therefore, HRTFs were measured on the artificial head VALDEMAR (developed at Aalborg University) with a directional resolution of 2°. The measurements were used to create HRTF data sets with low resolutions in the horizontal, frontal and median planes from which interpolations were done.

A listening experiment was done to determine when the differences between measured and interpolated HRTFs are audible. Sounds were filtered by either measured or interpolated HRTFs and presented through headphones in a three-alternative forced choice (3-AFC) test. Both static and dynamic sound sources were implemented.

2 Method

Clearly the directional resolution needed for measurements depends on the interpolation strategy. Here, the measured HRTFs were decomposed into minimum phase components (one for each ear) and a pure delay as interaural time difference (ITD). Interpolations were made by linear interpolation of the minimum phase components (in the time domain) after which the measured ITDs were reapplied. In this way only the audibility of errors in the minimum phase components were tested and additional errors due to a wrong ITD were avoided.

In the listening test for static sound sources the target HRTFs were selected at a regular spacing of 30°, giving 12 directions per plane. Due to the symmetry in the horizontal and frontal planes 7 directions were tested. Interpolations were made from measurements spaced 4°, 8°, 16°, 24°, 36° and 60° apart, with the measured 'target' HRTF in the middle of the interval. A 650 ms long pink noise signal with a 50 ms cosine fade-in and fade-out was filtered by either the measured or interpolated HRTFs.

The test was also done for dynamic movements through 90° arcs on the horizontal, frontal and median planes. The movements were centred directly in front, behind, left, right, above or below the listener. Interpolations were made from measurements spaced 8°, 12°, 24°, 36°, 60° and 90° apart to a resolution of 2°. A 1.5 s long pink noise signal with a 50 ms cosine fade-in and fade-out was used. For both static and dynamic sound sources the silences between the stimuli in a 3-AFC presentation were 80 ms.

The sounds were presented through a Beyerdynamic DT990 headphone, equalised from measurements made on the artificial head. The experiment was computer controlled and answers were collected by means of an electronic tablet.

A group of 8 people (4 male and 4 female) participated in the listening experiment. They had normal hearing and participated in previous experiments at the laboratory. Every listener received 468 presentations for static sound sources and 180 presentations for dynamic sound sources. The order of the presentations was randomised for every listener and divided into listening sessions (either static or dynamic) of 10-12 minutes.

The listener's task was to detect differences between the signals independent of the nature of the difference. The listener was allowed to repeat a presentation once, after which he/she had to select one of the options in the 3-AFC.

3 Results and discussion

The results are presented as the percentages of correct answers for every condition. Since there are 24 answers per condition the null hypothesis (that the listeners were guessing) is rejected if the number of correct answers is above 52% (5% significance level). Therefore, in Figures 1a-1c and Figures 2a-2c white fields indicate that the differences between measured HRTFs and interpolated HRTFs are undetectable.

A polar co-ordinate system is used with poles above and below the head. Horizontal planes are indexed by the azimuth angle, where 0° is in front, 90° is on the left, -90° is on the right and 180° is behind the head. Elevation angles are relative to the horizontal plane, i.e. 0° elevation is the horizontal plane, 90° is above and -90° is below the head.

3.1 *Static sound sources*

The results for static sound sources are presented in Figure 1a-1c. Notice for the horizontal and frontal planes that the data are the same on the left and right sides of the head.

For interpolations in the horizontal plane a resolution of 16° is not sufficient to avoid detection on the sides of the head, whereas lower resolutions are allowed in front and behind. In the frontal plane less than 8° resolution is needed for elevations of 0° and below whereas low resolutions are allowed above the head. Also in the median plane it is seen that resolutions of 24° are generally sufficient above the head. In front and behind the head, however, relatively high resolutions are required. It is observed that higher resolutions are required for interpolations in the vertical direction (in the frontal and median planes) than in the horizontal direction.

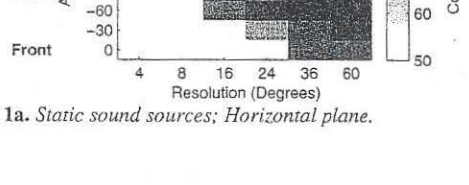

Figure 1a. *Static sound sources; Horizontal plane.*

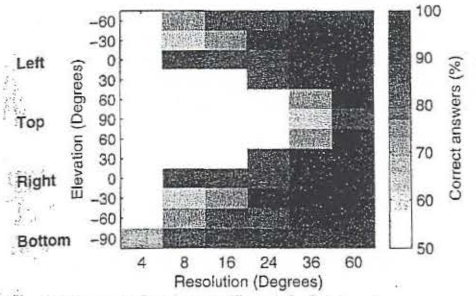

Figure 1b. *Static sound sources; Frontal plane.*

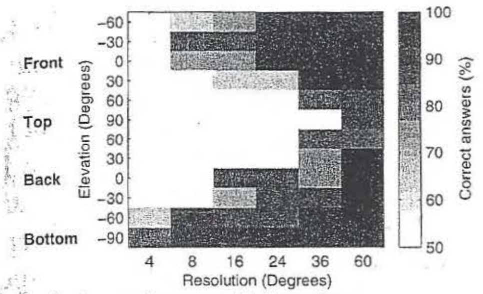

Figure 1c. *Static sound sources; Median plane.*

3.2 *Dynamic sound sources*

The results for the dynamic sound sources are given in Figures 2a-2c for the horizontal, frontal and median planes. Figures 2a-2c are different from Figures 1a-1c in the sense that the latter shows results for sound sources with a fixed direction whereas the former correspond to sounds moving on a path.

There is a striking resemblance between the static and dynamic results. However, the errors are much less detectable for dynamic sound sources. Notice, furthermore, that lower resolutions were used for the dynamic than the static sources. In the horizontal plane 36° resolution cannot be allowed on the sides whereas the interpolations in the front and back are much less critical. The highest resolution is needed below the head where 12° is not sufficient to avoid detection. When the sound source is in front a lower resolution is needed for interpolations in the vertical than the horizontal direction, as was the case for static sound sources.

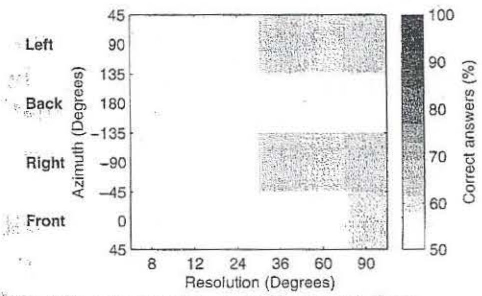

Figure 2a. *Dynamic sound sources; Horizontal plane.*

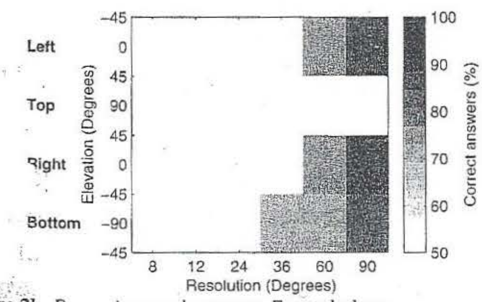

Figure 2b. *Dynamic sound sources; Frontal plane.*

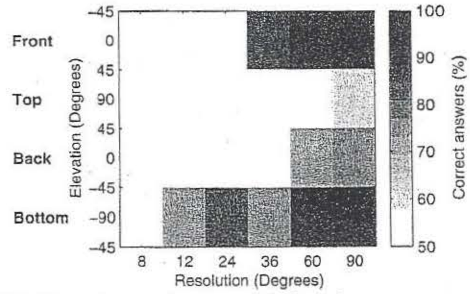

Figure 2c. *Dynamic sound sources; Median plane.*

4 Results applied to all directions

The HRTFs used for implementing static sound sources in the listening experiment were analysed and it was found that the audibility of interpolation errors can be predicted by a simple spectral distance measure. The error between a measured and interpolated HRTF is calculated as the difference in magnitude (in dB) on a logarithmic scale between 100 Hz and 20 kHz. The absolute errors of the left and right HRTFs are added and if the mean (across frequency) is larger than 1 dB the difference will be audible.

HRTFs were measured on VALDEMAR with a resolution of 2° on the whole sphere around the head. The criteria for audibility was applied to these HRTFs in order to extrapolate the results of the listening experiment to all directions. Figure 3 shows the resolution needed to avoid detection for vertical (worst case) interpolations. The mean of directions on the left and right sides of the head is shown since the results were found to be nearly symmetrical.

In the region above 60° elevation a resolution of 24° is sufficient to avoid detection. Above 45° elevation a resolution of 16° or higher may be used. Behind the head, for azimuth angles larger than 120° and elevation angles between -60° and 45°, a resolution of 8° or higher is allowed, whereas for the remaining directions a resolution of at least 4° is generally required to avoid detection.

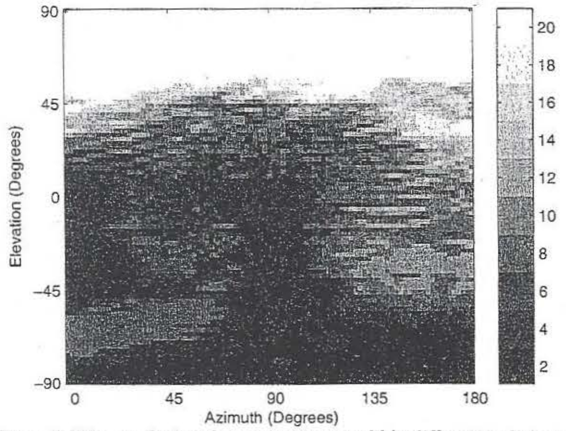

Figure 3. *The resolution that ensures no audible differences between measured and interpolated HRTFs.*

5 Acknowledgement

Many thanks to Bjarke Phil Bovbjerg, Jacob Florian Randbæk and Claus Vestergaard Skipper. This work was partly funded by the Danish National Centre for IT Research.

6 References

[1] F. Christensen, H. Møller, P. Minnaar, J. Plogsties, S. K. Olesen, "Interpolating between head-related transfer functions measured with low directional resolution", *Proceedings of the 107th Convention of the Audio Engineering Society*, New York, USA, 1999 (preprint 5047).